

atoms in the subgraph. This would represent a major improvement. Unfortunately, the algorithm is now NP-complete with respect to sequential processors.

We proposed a "star" configuration architecture; a central processor with a communication line to a number of lower processors, with no direct communication allowed between the lower processors. Each processor has a small amount of memory as a "working space", which avoids the problems inherent in shared memory. The communication packets are compact to reduce the storage and communication burden on the central processor.

The simulation algorithm called MOLSIM was implemented using the SIMULA language. We studied this algorithm as a function of the number of processors and the nature of the particular graphs being matched (6-31 non-H atoms). We found an average utilization of 84% on a 25 cpu machine (figured as total processor time/real time), but only 60% on a 50 cpu machine although for some structure matching questions, the efficiency reached 97%. The average speed enhancement using this size machine (50 processors) was a factor of 30. A real machine of the architecture needed to run this algorithm exists at Purdue University, and time is being requested to test the algorithm in real time. This algorithm is a unique approach to the problem of graph matching and will likely become practical when parallel processors are commonly available and inexpensive. (This work is submitted for publication in J. Chem. Inf. Comput. Sci.) *C.1.b MCS-- Maximal Common Subgraph Search.*

The second area of starting material strategy work this year has been in solving the class IV problem given above. Our solution to this problem involves development of a new efficient maximal common subgraph matching algorithm. Since chemists represent organic molecules as graphs, computational chemists need graph theoretical techniques such as graph isomorphism, subgraph search, and maximal common subgraph search. Of these three important procedures, maximal common subgraph search (MCSS) remains the most difficult and least utilized.

The extensive computational demands of MCSS has restricted its possible uses. We have previously noted that maximal common subgraph search could be useful in our starting material selection program, SST, but that the computational demands were too rigorous.\* Cone et al. who has used MCSS in their "self-training interpretive and retrieval system" (STIRS), has noted that other potential uses of for MCSS include computer-assisted organic synthesis and structure activity studies.\*\*

Given two graphs, finding a *common subgraph* involves discovering the assignment of some of the nodes and edges of one graph onto the other graph while preserving the adjacency relationships of the nodes. The *size* of the common subgraph is the number of edges preserved in the assignment. If there exists no common subgraph of larger size, the common subgraph is called *maximal*.

Our approach to MCSS was to reduce redundant searching and try to shrink the size of the search space. We observe that most libraries of chemicals have compounds which have similarities; by capitalizing on these similarities we might be able to reduce the search space. The essential principle is that if we know a relationship between library graph A and B we can relate query graph C to A, then we may therefore already know something about the relationship of C to B. We establish the relationship between A and B in a one-time-only preprocessing of the library.

---

\* W.T. Wipke, D. Rogers, *J. Chem. Inf. Comput. Sci.*, 1984, (in press)

\*\* M.M. Cone, R. Venkataraghavan, F.W. McLafferty, *J. Am. Chem. Soc.*, 99, 7888, (1977)

The objective of our study was to demonstrate the feasibility and study properties of such an algorithm. The following design features were deemed important.

- The processed library should be storage efficient.
- It should be possible to add compounds *incrementally* to the file, to avoid the cost of reprocessing the file whenever changed.
- The time required to store a new compound in the processed library should be minimal; preferably, an upper limit on this time requirement should be known.
- It should be possible to create the compound from its processed form, so that both processed and non-processed libraries do not have to be kept.

We wished the system to be useful for a range of non-preselected queries, therefore we ruled out simply "training" the system for a small class of query choices or depending on detailed knowledge of the allowed queries.

We had noticed that identical common subgraph candidates are often generated during the search of different compounds against a query. We envisioned collecting these common segments into an intermediate graph, to allow the search to be conducted once over the common feature. If our search procedure can take advantage of the abstraction of the common graph from two or more test compounds, then the number of attempted subgraph matches will be reduced.

A non-recursive FORTRAN algorithm was implemented for the MCS-1 program using a tree-structured storage file. The tree storage search algorithm gives reductions from 70 to 90% in the search relative to conventional unstructured sequential storage systems. Reductions are especially good for the case where the library contains smaller graphs or a series of similar graphs. The general structure of the search tree is determined early in its creation; once the major nodes of the tree are established, addition of compounds rarely alters it significantly. Sorting experiments showed that the tree can be "seeded" in such a way that improved results can be obtained when searching over the seeded library relative to the unseeded library.

We have applied this MCS-1 algorithm to the Class IV starting material recognition problem. The initial abstraction of the starting material library resulted in an abstracted library of significantly reduced size. Organizing this abstracted library in a tree-structured form allows the discovery of starting materials which do not have a subgraph or super graph relationship with the target. A trial run with morphine was successful in pointing out an interesting starting material candidate not found by the SST program. This work is being submitted to the J. Chem. Inf. Comput. Sci..

### *C.2 XENO Program Developments*

The metabolic fate of various compounds in the human body is extremely complex, yet extremely important for it is known that through metabolism certain otherwise harmless compounds are converted into toxic and possibly carcinogenic agents. Because of this complexity it is difficult, looking at a given compound, to forecast potential biological activity of that given compound. The objective of this proposal is to develop a practical computer program by which a biochemist or metabolism expert can explore the metabolites of a given compound and be alerted to the plausible biological activity of each metabolite.

This research aims to explore the degree to which current knowledge of metabolism can be used by a computer program to make reasonable projections of what metabolites

might result from exposure of a compound to a biological system. The project involves representing in a computer the metabolic processes with all known specificities and applying these processes to a xenobiotic compound to generate a set of plausible metabolites which may themselves be further metabolized. We also plan an evaluation module to appraise the plausible biological activities of each metabolite using a rule base to relate chemical structure to biological activity. Thus the attention of the experimentalist will be attracted to those metabolites that are likely to be biologically active.

### C.2.a Atomic Charges.

During the past year, we implemented rapid atomic charge calculation algorithms to be used with the XENO program in order to better predict the biological activity of metabolites. The algorithms were based on Gasteiger's PEOE (partial equalization of orbital electronegativity)\* and SD-POE (sigma dependent pi orbital electronegativity)\*\* models. The PEOE model has been used for sigma charge calculations for sigma bonded and non-conjugated pi systems. The SD-POE model has been used for conjugated aliphatic and single ring aromatic molecules. For polyaromatic systems, the pi charge calculations are being implemented using the SD-POE model and the Longuet-Higgins approximation\*\*\*. In the PEOE model, the SD-POE model, and polyaromatic hydrocarbon pi charge calculation, atoms are characterized by their orbital electronegativities.

We have shown that the charges so calculated are reasonable when compared to the work of others in the literature. Our purpose is then to correlate the biological activity of metabolites with the atomic charge on electrophilic centers in the metabolites. We also think that metabolism itself can be controlled to some extent by atomic charges so the metabolic transforms may make use of this data eventually.

### C.2.b pKa Calculations.

We have continued work on the problem of the estimation of the dissociation constants for organic acid and bases to be used to increase the expertise of the XENO program. We have been investigating two approaches to do this type of estimation: LFER (linear free energy relations), and theoretical or quantum chemical approaches\*\*\*\*.

The LFER computation is performed automatically by first selecting appropriate skeleton structures from a library, then recognizing attached groups and finally calculating the pKa from the relevant equations and group substituent constants. This is the first automatic pKa estimation algorithm ever developed and promises wide utility on its own outside of the XENO program.

To determine the most representative pKa if more than one acid or base center is present, several empirical rules are followed:

- **acids** - use the ionized form of the acid center with the lowest pKa as a substituent for the pKa calculation of the next lowest acid, and so on.

---

\* J. Gasteiger and M. Marsili, *Tetrahedron lett.*, **34**, 3181, (1981)

\*\* J. Gasteiger and M. Marsili, *Tetrahedron*, **38**, 3219, (1980)

\*\*\* H. C. Longuet-Higgins, *J. Chem. Phys.*, **18**, 275, (1950)

\*\*\*\* D.D. Perrin, B. Dempsey, and E.P. Serjeant, "pKa Prediction for Organic Acids and Bases", Chapman and Hall, New York, 1981.

- **bases** - use the protonated form of the base center with the highest pKa as a substituent for the pKa calculation of the next highest base, and so on.
- **strong acids and bases** - compute pKa's for the acids first, then use ionized form of the acids as substituents to compute pKa's of the bases.
- **strong acids and weak basis** - compute pKa's for bases first, then use protonated form of the bases as substituents to compute pKa's of the acids.
- **weak acids and weak basis** - compute pKa's for weak bases first, then use deprotonated form of bases as substituents to compute pKa's for the acids.

*C.2.g Collaborative Efforts.* In the past year most work was aimed at program development rather than application to laboratory problems, however in the next year we do expect after completion of the current program modifications to perform several practical analyses in conjunction with our collaborators at ICI of UK, NIH, and other parties that have indicated interest.

The SECS project continues to have collaborations with the pharmaceutical industry who are adding chemical transforms and doing some joint program development, for example, Dr. Yanaka continued work started at Santa Cruz after he returned to Kureha Chemical in Japan and a paper has been prepared on that work.

#### *D. List of Current Project Publications*

1. Wipke, W.T., and Rogers, D.: *Rapid Subgraph Search Using Parallelism* J. Chem. Inf. Comput. Sci (submitted 24 April 84)
2. Wipke, W.T.: "An Integrated System for Drug Design" in *Computers A-Z: A Manufacturer's Guide to Hardware and Software for the Pharmaceutical Industry* Aster Publishing Co., Springfield, Oregon, (in press)
3. Wipke, W.T. and Huber, M.: *Symmetry and organic synthetic design.* Accepted in Tetrahedron.
4. Wipke, W.T., Ouchi, G.I. and Chou, J.T.: *Computer-Assisted Prediction of Metabolism.* IN L. Goldberg (Ed.), STRUCTURE-ACTIVITY CORRELATIONS AS A PREDICTIVE TOOL IN TOXICOLOGY. Hemisphere Publishing Corp., New York, 1983.
5. Johnson, C.K., Thiessen, W.E., Burnett, M.N., Condran, P. Roulan, A., Yanaka, M. and Wipke, W.T.: *Systematic derivation of chemical procedures for transforming surplus hazardous chemicals to useful products, J. of Hazardous Materials.* (In press)
6. Dolata, D.P.: *QED: Automated Inference in Planning Organic Synthesis* (Ph.D. dissertation). University of California, Santa Cruz, 1984.
7. Rogers, D.: *Artificial Intelligence in Organic Chemistry. SST: Starting Material Selection Strategies* (Ph. D. dissertation). University of California, Santa Cruz, 1984.

Wipke, W.T., and Rogers, D.: *Artificial Intelligence in Organic Synthesis. SST: Starting Material Selection Strategies. An Application of Superstructure Search.* J. Chem. Inf. Comput. Sci., 24:0000, 1984.

#### *E. Funding Status*

1. Computer-Assisted Prediction of Xenobiotic Metabolism  
Principal Investigator: W. Todd Wipke, Professor, UCSC  
Agency: NIH, Environmental Health Sciences; No: ES02845-02  
4/1/82-3/31/85 \$257,801 TDC  
4/1/84-3/31/85 \$ 89,140 TDC
2. Graphical Display of Chemical Inferences and Molecular Relationships  
Principal Investigator: W. Todd Wipke, Professor, UCSC  
Agency: Evans and Sutherland Corporation  
Gift of PS300 B/W High Performance Graphical Display System  
Permanent, value \$95,000 TDC
3. Computer Synthesis  
Principal Investigator: W. Todd Wipke, Professor, UCSC  
Agency: Stauffer Chemical Company  
Permanent, \$6,000 TDC

#### *F. Research Environment*

At the University of California, Santa Cruz, we are connected to the SUMEX-AIM resource by a 4800 baud multiplexed leased line. Our video terminals consist of a Z-29, DM-3025, TI745, CDI-1030, DIABLO 1620, and an ADM-3A. We have a PS300 graphic display which is driven by SUMEX. UCSC has only a small IBM 370/145, a PDP-11/45, 11/70 and a VAX 11/780, (the 11's are restricted to running small jobs for student time-sharing) all of which are unsuitable for our current research. The SECS laboratory is located in 125 Thimann Laboratories, adjacent to the synthetic organic laboratories at Santa Cruz.

## **II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE**

### *A. Medical Collaborations and Program Dissemination via SUMEX*

SECS is available in the GUEST area of SUMEX for casual users, and in the SECS DEMO area for serious collaborators who plan to use a significant amount of time and need to save the synthesis tree generated. Much of the access by others has been through the graphic terminal equipment at Santa Cruz, so much more convenient for structure input and output. Demonstrations and sample synthetic analyses were generated for numerous visitors from the US and abroad. Demonstrations of SECS in Sweden were performed by Dr. R. E. Carter, University of Lund, Sweden, at many universities and companies.

Professor Wipke has also used several SUMEX programs such as CONGEN in his course on Computers and Information Processing in Chemistry. Communication between SECS collaborators is facilitated by using SUMEX message drops, especially when time differences between the U.S. and Europe and Australia makes normal telephone communication difficult. Testing and collaboration on the XENO and FSECS project with researchers at the NCI depend on having access through SUMEX and TYMNET.

Collaboration with Lund University. The introduction of SECS to organic chemists in Sweden was one of the seeds that led to the establishment of a computer graphics laboratory for organic chemistry at the University of Lund, with strong support from a government agency, the National Swedish Board for Technical Development (STU).

Interest in the applications of computers and computer graphics in organic chemistry has spread very rapidly throughout the country, and chemists at all of the major Swedish universities, as well as in the pharmaceutical industry, have taken steps to participate in the exciting developments in this field.

Interest in the pedagogical value of SECS at the graduate level has led to its use to illustrate the concepts of retrosynthetic planning and analysis in conjunction with a course given by Prof. Paul Helquist (SUNY, Stony Brook): "Synthetic Organic Chemistry---Modern Methods and Strategy". The same course was given at the Royal Institute of Technology in Stockholm, and at the Chemical Center in Lund, using SECS as an integral part of the course.

A Workshop on Computers in Organic Chemistry was sponsored by STU on May 17-18, 1983, in Gothenburg to help organic chemists in Sweden enter this area of work. Daniel P. Dolata from Prof Wipke's group in SC was an invited speaker.

A chemist from Lund, Dr. Alvin Ronlan, spent a sabbatical leave with the Wipke group, and a graduate of the Wipke group, Dr. Dolata, is spending a postdoctoral stay in Lund.

In collaboration with the SECS group at UCSC, Dolata will install the SECS program and the QED program on a new VAX 11/780 for the use of the chemists in Lund, and will continue research with QED. For example, it would be of interest to develop rule bases to assist the chemist in structure elucidation, and structure-activity relationships.

Another area of collaboration involves compilation of chemical transforms by the chemists in Lund. Some of the chemists in Lund work with natural products (isolation and synthesis), with a view toward the discovery and characterization of physiologically active substances. For example a strongly mutagenic compound has been isolated from a Swedish mushroom (*Lactarius vellereus*), its structure determined, and a total synthesis elaborated. In other work, a traditional abortifacient from Bangladesh is being isolated from plant material, and a psychoactive substance is being extracted from the leaves of a Nigerian plant. A collaboration with a university in Holland is now developing along similar lines, and Cornell University is planning a similar center for computer applications in chemistry.

#### *B. Examples of Cross-fertilization with other SUMEX-AIM Projects*

The AILIST bulletin board has been used extensively for interacting with many projects and locating references for further information related to program design and AI technology. There are no longer any other chemical or biochemical projects on SUMEX so our interaction with the community is limited to AI technology interchange, attending seminars at Stanford, etc.

#### *C. Critique of Resource Services*

SUMEX-AIM gives us at UCSC, a small university, the advantages of a larger group of colleagues, and interaction with scientists all over the country. Previously we were provided very good service by SUMEX-AIM, but since 1 April 1984, the computer service has been very poor. Although the National AIM usage of SUMEX has been small, our project has been put in a separate class with a 3% cpu limitation. This is a very severe restriction which prevents short usage peaks from being averaged with other users. Our project is the only project subjected to such limitations. The poor response time we are observing (load averages of 25-50!) is significantly hindering our ability to perform the research NIH funded to be done on SUMEX. This is worsened by the fact we are in the last year of the project.

#### *D. Collaborations and Medical Use of Programs via Computers other than SUMEX*

SECS 2.9 has been installed on the CompuServe computer networks for the past three years so anyone can access it without having to convert code for their machine. This has proved very useful as a method of getting people to experiment with this new technology. Dr. George Purvis of Battelle has accessed SECS via CompuServe, as has Gene Dougherty of Rohm and Haas and many others. SECS also resides on the Medicindat machine at the University of Gothenborg, Sweden, and is available all over Sweden by phone. Similarly in Australia, SECS resides at the University of Western Australia and is available throughout Australia over CSIRONET. A lecture series was given on SECS in Tokyo and SECS has been installed at two locations in Japan. FSECS has been installed on a DEC-10 at Oak Ridge National Laboratory and serves for collaborative development of that approach with Carroll Johnson. PRXBLD has been disseminated to over 30 sites on various types of computers including DEC-10, DEC-20, IBM, VAX, PRIME and Honeywell.

### **III. RESEARCH PLANS (4/84-4/85)**

*A. Near-Term Project Goals and Plans* Our research projects will move off SUMEX-AIM by 31 March 1985 to some other as yet unspecified computer system. Therefore our research objectives on SUMEX-AIM are to complete research in progress, consolidate programs and files for moving to another system. The QED and SST projects have been completed and the first phase of the RXAN project outlined last year has been completed. On the SECS project, the reaction library is being extended by Dr. Iwataki. We will continue to collaborate with coworkers in SECS research on other machines but on SUMEX will primarily be preparing SECS for removal from SUMEX. We are exploring PROLOG as a replacement for the QED system and plan some preliminary sample PROLOG programs to compare the capabilities of PROLOG and QED. But the majority of our activities will be aimed at completion of the XENO program.

*XENO Goals* Our objectives for this year follow our plan in the original proposal. Basically in this next year we plan to complete the implementation of algorithms not yet completed and focus on testing with applications to demonstrate the current power of XENO on typical laboratory metabolism problems. In this last year of this project our goal is to bring XENO to a relatively stable finished point which will be useful to other researchers. We believe we will complete the algorithms in progress, document them and submit publications on all of the work within the year. The major areas of focus are listed below.

*A.1 Atomic Charge Calculations.* We also plan to complete our correlations between atomic charges and sites of metabolism, as we have already done with bond reactivity. When such correlations are established, then they can be used to guide XENO to apply metabolic transforms more selectively to the most active parts of the molecules.

*A.2 pKa Calculations.* We plan to complete testing of the pKa algorithm on different groups on metabolites so that information may be used by XENO for activity evaluation, selection of further possible metabolism, and estimation of excretion and transport.

*A.3 Three-Dimensional Criteria.* We plan to complete our work on three-dimensional constraints that apply to metabolism which have been obtained through study of many metabolic studies in the literature. This will require extensions of the

ALCHEM language to accommodate new types of three-dimensional relationships such as overall molecular size, width, thickness, ratio of length to width, etc.

**A.4 Log P Calculation.** In order to more accurately estimate the possibility of excretion and binding for metabolism, we plan to incorporate a log P calculation module. This will provide the partition coefficient between octanol and water. There are already programs to calculate log P, and these have been shown to be very accurate. Our objective, time permitting, will be to include such a module in the XENO program and correlate log P with metabolism transform application and with excretion.

#### ***B. Justification and Requirements for Continued Use of SUMEX***

The XENO project which resides on SUMEX is in its last year of support, consequently we need to complete that research on the SUMEX machine. By 31 March 1985, we plan to move XENO and all our research off SUMEX onto some other computer. We are currently exploring what machine may be suitable and available. After 1 April 1985 we will not need SUMEX for computational support, but will need access to be able to retrieve certain files from archive, respond to electronic mail, and continue to participate in the AIM scientific interchange through electronic mail and bulletin boards. It is not practical to retrieve every file we have ever archived, it would use too much SUMEX operator time, and it is unnecessary as long as we can access them if we need them in the future. That access would not require significant resources.

However prior to 31 March 1985 we have obligations to complete the research on the XENO project supported by NIH and need sufficient SUMEX cpu time to accomplish this goal. This means normal editing, compile, load, and test executions plus some application runs to some metabolic problems. It appears the current removal of our project from the National AIM portion of the SUMEX-AIM resource and placement in a class restricted to 3% peak utilization is hindering the research productivity of this project. We are experiencing load averages of 25-50 a high percentage of the time. We request to have our project placed back in the National AIM portion of the SUMEX-AIM resource as we were allocated, and we will carefully monitor to see that our resource utilization does not exceed our quota of time. We feel this is a reasonable request in light of the mission of SUMEX-AIM to the National community of which this project is a part.

**C. Needs Beyond SUMEX-AIM** As mentioned above our project needs additional computing resources and we are exploring acquiring a computer for installation at UCSC and obtaining the necessary resources to support it. We are seeking information about comparisons between machines and cost effectiveness of different hardware combinations.

#### ***D. Recommendations for Community and Resource Development***

It appears the SUMEX-AIM resource is increasingly becoming basically a Stanford resource and that there is a difference between the portion of the resource allocated to the National community and the portion actually used by the National community. Our project is part of the National community and in need of better service, we hope that can be improved.

An important part of medicine is treatment of diseases with drugs, chemicals, chemicals that were designed and synthesized by chemists. Since the termination of the DENDRAL project, there seems to be declining support for artificial intelligence applications in chemistry. We feel that support of this area is essential to the advancement of medicine in this country. The lack of chemists on NIH Research Resources computing peer review is contributing to the problem. In general the AIM community would benefit by involving disciplines other than computer science.

## **II.A.2.5. SOLVER Project**

### **SOLVER: Problem Solving Expertise**

**Dr. P. E. Johnson**  
**Center for Research in Human Learning**  
**University of Minnesota**

**Dr. W. B. Thompson**  
**Department of Computer Science**  
**University of Minnesota**

## **I. SUMMARY OF RESEARCH PROGRAM**

### *A. Project Rationale*

This project focuses upon the development of strategies for discovering and documenting the knowledge and skill of expert problem solvers. In the last fifteen years, considerable progress has been made in synthesizing the expertise required for solving extremely complex problems. Computer programs exist with competency comparable to human experts in diverse areas ranging from the analysis of mass spectrograms and nuclear magnetic resonance (Dendral) to the diagnosis of certain infectious diseases (Mycin).

Design of an expert system for a particular task domain usually involves the interaction of two distinct groups of individuals, "knowledge engineers," who are primarily concerned with the specification and implementation of formal problem solving techniques, and "experts" (in the relevant problem area) who provide factual and heuristic information of use for the problem solving task under consideration. Typically the knowledge engineer consults with one or more experts and decides on a particular representational structure and inference strategy. Next, "units" of factual information are specified. That is, properties of the problem domain are decomposed into a set of manageable elements suitable for processing by the inference operations. Once this organization has been established, major efforts are required to refine representations and acquire factual knowledge organized in an appropriate form. Substantial research problems exist in developing more effective representations, improving the inference process, and in finding better means of acquiring information from either experts or the problem area itself.

Programs currently exist for empirical investigation of some of these questions for a particular problem domain (e.g. AGE, UNITS, RLL). These tools allow the investigation of alternate organizations, inference strategies, and rule bases in an efficient manner. What is still lacking, however, is a theoretical framework capable of reducing dependence on the expert's intuition or on near exhaustive testing of possible organizations. Despite their successes, there seems to be a consensus that expert systems could be better than they are. Most expert systems embody only the limited amount of expertise that individuals are able to report in a particular, constrained language (e.g. production rules). If current systems are approximately as good as human experts, given that they represent only a portion of what individual human experts know, then improvement in the "knowledge capturing" process should lead to systems with considerably better performance.

### *B. Medical Relevance and Collaboration*

Collaboration with Dr. James Moller MD in the Department of Pediatrics, Dr. Donald Connelly MD in the Department of Laboratory Medicine, at the University of Minnesota. Collaboration with Dr. Eugene Rich MD and Dr. Terry Crowson MD at St. Paul Ramsey Medical Center.

### *C. Highlights of Research Progress*

Accomplishments of This Past Year -- Prior research at Minnesota on expertise in diagnosis of congenital heart disease has resulted in a theory of diagnosis and an embodiment of that theory in the form of a computer simulation model, *Galen*, which diagnoses cases of congenital heart disease [Thompson, Johnson & Moen, 1983].

Galen is descended from two earlier programs written here at Minnesota: *Diagnoser* and *Deducer* [Swanson, 1977]. *Deducer* is a program that builds hemodynamic models of the circulatory system that describe specific diseases. The models are built by using knowledge about how idealized parts of the circulatory system are causally related. *Diagnoser* is a recognition-driven program that performs diagnoses by successively hypothesizing one or more of these models and matching them against patient data. The models that match best are used as the final diagnosis. A series of experiments carried out at Minnesota have shown that *Diagnoser/Deducer* performs as well (and sometimes better) than expert human cardiologists [Johnson et al., 1981].

Despite their early successes, *Diagnoser* and *Deducer* did not have a clear, comprehensible structure that is required for the kind of experiments we wish to perform. *Galen* was built to remedy this problem, taking advantage of the experience gained in the design of *Diagnoser* and *Deducer*.

*Galen* consists of four major components: a working memory called the *scratchpad*, a knowledge base of rules and hypotheses, a procedure called the *proposer* and a procedure called the *reviewer*.

The *scratchpad* contains data about the problem that *Galen* is trying to solve and the hypotheses that are being investigated to explain that data. In effect, the *scratchpad* represents *Galen's* current execution state.

Rules are pattern-action pairs. The pattern part of a rule describes a possible state of the *scratchpad*. Patterns can contain imbedded logical connectives (e.g. ANDs, ORs, NOTs) and can be constructed to match at varying levels of detail. The action part is a procedure that is executed if the pattern part matches the *scratchpad's* contents. Each action part writes an assertion on the *scratchpad* about a hypothesis, together with the evidence for making that assertion. These assertions can express that a new hypothesis is being considered, or that an old hypothesis has been accepted, rejected, confirmed or disconfirmed. Action parts can also assert that a hypothesis is sufficient to solve the current problem, or that the problem is not solvable.

Because the pattern parts of rules can examine anything on the *scratchpad*, it is possible to express rules about hypotheses as well as rules about problem data. In particular, this makes it possible to directly examine the accumulated evidence for and against each currently contending hypothesis, making numerical measures of certainty unnecessary.

A hypothesis is simply a named collection of rules. The hypotheses in *Galen's* knowledge base can be thought of as a directed graph, in which vertices are hypotheses and edges are rules. One hypothesis "points to" another if the first hypothesis contains a rule whose action part can assert something about the second.

The level of detail of such a knowledge base leads to serious problems with the computational complexity of search processes. Galen focuses its computational resources so that the knowledge embodied in the graph of hypotheses can be used in an efficient manner. Successful diagnoses result from good first hypotheses about possible defects and efficient mechanisms for refining these hypotheses.

Galen works by using the proposer and the reviewer to investigate hypotheses (i.e. search the graph) by applying rules (i.e. following the edges from one vertex to another). Whenever a new piece of problem data is written on the scratchpad, the proposer applies all relevant rules specific to the type of that piece of data. These rules write assertions on the scratchpad about new hypotheses, effectively identifying vertices in the graph that are worthy starting points for further search. Next, the reviewer applies all relevant rules contained in the hypotheses that are named in assertions on the scratchpad. Successfully applying one of these rules corresponds to propagating the search along a specific edge of the graph. The search is constrained because (1) only the most promising vertices in the graph are ever used to initiate search; (2) only a small number of edges are ever followed; and (3) most rules in a hypothesis deal with evidence for and against the hypothesis itself, giving a graph where the number of effectively outward-pointing edges at each vertex is small.

The read-propose-review cycle repeats in this way until some hypothesis has been shown correct, until the problem has been shown unsolvable, or until all the data has been examined.

Currently, data given to Galen is taken from a (possibly imaginary) patient's medical chart. Hypotheses in the knowledge base represent the ten most commonly occurring congenital heart diseases and their variants, useful intermediate physiological findings, and classes of hypotheses. Since hypotheses are implemented as named teams of production rules, it is also possible to represent other kinds of hypotheses should the need arise. Moreover, Galen has been constructed so that its inference engine does not contain any procedures specialized for pediatric cardiology. It is therefore conceivable to extend Galen to other domains if effective knowledge bases for those domains can be constructed.

To determine the generality of our model of expertise in diagnostic reasoning, we are also investigating domains outside medicine. As part of this effort, we have developed a computational model of the fault localization process in program debugging [Sedlmeyer, 1983] that is not based directly on Galen. As with our work in congenital heart disease, we have concentrated on the design of mechanisms for structuring problem specific knowledge and for focusing limited computational resources.

Research in Progress -- Since human experts are notoriously poor at describing their own knowledge, our work requires the creation of problem solving tasks through which experts can reveal criteria for initiating specific hypotheses and methods for investigating those hypotheses.

Current techniques of representing hypotheses and their expectations for diagnosis do not, however, provide much detailed information about the control processes experts use to guide their reasoning. Such control processes typically incorporate highly refined heuristics about which the experts are almost wholly unaware. To discover the needed control knowledge, we ask experts to complete tasks in which we have systematically perturbed aspects of the problem data. The data in these tasks are chosen so that members of an overlapping set of hypotheses will be suggested during while solving the problem. Success in solving such problems depends on the ability to overcome an initially plausible incorrect hypothesis in favor of a later, more correct alternative.

Several examples illustrate our approach. We are studying performance of Galen on "garden path" cases [Johnson & Thompson, 1981] that were initially misdiagnosed in hospital files. Analysis of such cases suggest that errors are made because experts rely on very efficient heuristics that are not universally correct. In one such example, a seemingly plausible hypothesis is suggested early in the case. Although the hypothesis superficially seems to explain what is observed about the patient, the hypothesis is incorrect. Because the incorrect hypothesis seems marginally adequate, it acts to prevent a more correct hypothesis from being suggested in its place. Success in such a case hinges on the ability to use the proper set of competing hypotheses in order to provide more than one explanation of the case data. Investigation of this phenomenon in human experts has suggested implementation of "transition rules" linking disease hypotheses in Galen. It has also suggested implementation of "monitor hypotheses" that watch for potential garden path errors and avoid them before they become serious.

We are also investigating several research questions relevant to the architecture of Galen. We have designed an interface to Galen so that users who are unfamiliar with the inner workings of the program can interactively enter case data. Designing the interface raised questions about what forms of data are necessary to adequately and completely represent all possible cases. We are also studying ways in which a causal reasoning component can be integrated with the prototypical reasoning components (the Proposer and Reviewer) that are already present in Galen. In particular, we are interested in studying ways in which causal reasoning can aid or replace prototypic reasoning when it becomes inadequate to reach a diagnosis.

In another project, we are investigating methods of probabilistic reasoning. Most systems rely on numerical schemes for weighting evidence or ranking observed data. These weights are often probabilistic in nature, but other schemes have also been used. Mycin, for example, uses certainty factors and PIP uses likelihoods composed of matching scores and binding scores. In contrast, humans do not seem to rely upon such numerical techniques. Research has shown that people are often quite poor at probabilistic reasoning. However, experts make decisions which involve weighting evidence and selecting from competing alternatives. They must utilize a reasoning process which serves as an alternative to a numerical weighting technique.

We believe the process of weighting alternatives along various criterial dimensions is not a domain specific technique, but rather a general process which is applied in specific instances. In the coming year, we will examine this process in various domains and attempt to utilize the results in designing more powerful reasoning techniques.

In the area of law, our work has focused on the area of corporate law (the problem of structuring a proposed corporate acquisition). We have collected data from 24 practicing lawyers and in the coming year a PhD thesis will be completed describing this work. In the coming year we will be also completing a study of the corporate acquisition problem in management in order to further refine our knowledge capturing tools.

#### *D. List of Relevant Publications*

1. Connelly, D. and Johnson, P.E.: *Medical problem solving*. Human Pathology, 11(5):412-419, 1980.
2. Elstein, A., Gorry, A., Johnson, P. and Kassirer, J.: *Proposed Research Efforts*. IN D.C. Connelly, E. Benson and D. Burke (Eds.), CLINICAL DECISION MAKING AND LABORATORY USE. University of Minnesota Press, 1982, pp. 327-334.

3. Feltovich, P.J.: *Knowledge based components of expertise in medical diagnosis*. Learning Research and Development Center Technical Report PDS-2, University of Pittsburgh, September, 1981.
4. Feltovich, P.J., Johnson, P.E., Moller, J.H. and Swanson, D.B.: *The Role and Development of Medical Knowledge in Diagnostic Expertise*. IN W. Clancey and E.H. Shortliffe (Eds.), READINGS IN MEDICAL AI. (In press)
5. Johnson, P.E.: *Cognitive Models of Medical Problem Solvers*. IN D.C. Connelly, E. Benson, D. Burke (Eds.), CLINICAL DECISION MAKING AND LABORATORY USE. University of Minnesota Press, 1982, pp. 39-51.
6. Johnson, P.E.: *What kind of expert should a system be?* J. Medicine and Philosophy, 8:77-97, 1983.
7. Johnson, P.E., *The Expert Mind: A new Challenge for the Information Scientist* IN Th. M. A. Bemelmans (Ed.), INFORMATION SYSTEM DEVELOPMENT FOR ORGANIZATIONAL EFFECTIVENESS, Elsevier Science Publishers B. V. (North-Holland), 1984.
8. Johnson, P.E., Severance, D.G. and Feltovich, P.J.: *Design of decision support systems in medicine: Rationale and principles from the analysis of physician expertise*. Proc. Twelfth Hawaii International Conference on System Science, Western Periodicals Co. 3:105-118, 1979.
9. Johnson, P.E., Duran, A., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P. and Swanson, D.: *Expertise and error in diagnostic reasoning*. Cognitive Science 5:235-283, 1981.
10. Johnson, P.E. and Hassebrock, F.: *Validating Computer Simulation Models of Expert Reasoning*. IN R. Trappl (Ed.), CYBERNETICS AND SYSTEMS RESEARCH. North-Holland Publishing Co., 1982.
11. Johnson, P.E. and Thompson, W.B.: *Strolling down the garden path: Detection and recovery from error in expert problem solving*. Proc. Seventh IJCAI, Vancouver, B.C., August, 1981, pp. 214-217.
12. Johnson, P.E., Hassebrock, F. and Moller, J.H.: *Multimethod study of clinical judgement*. Organizational Behavior and Human Performance 30:201-230, 1982.
13. Moller, J.H., Bass, G.M., Jr. and Johnson, P.E.: *New techniques in the construction of patient management problems*. Medical Education 15:150-153, 1981.
14. Swanson, D.B.: *Computer simulation of expert problem solving in medical diagnosis*. Unpublished Ph.D. dissertation, University of Minnesota, 1978.
15. Swanson, D.B., Feltovich, P.J. and Johnson, P.E.: *Psychological Analysis of Physician Expertise: Implications for The Design of Decision Support Systems*. In D.B. Shires and H. Wold (Eds.), MEDINFO77, North-Holland Publishing Co., Amsterdam, 1977, pp. 161-164.
16. Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning*. Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

17. Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: *Knowledge-based fault localization in debugging*. The Journal of Systems and Software (in press).
18. Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: *Diagnostic reasoning in software fault localization*. Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.
19. Sedlmeyer, R.L., Thompson, W.B., & Johnson, P.E.: *Knowledge-Based Fault Localization in Debugging*. The Journal of Systems and Software (in press).

#### *E. Funding and Support*

Work on the SOLVER project is currently supported by a grant from the Control Data Corporation to Paul Johnson (\$90,000; 1983-85) and by a grant from the Microelectronics and Information Sciences Center at the University of Minnesota to Paul Johnson, William Thompson and two colleagues (\$800,000; 1984-87).

## **II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE**

### *A. Medical Collaborations and Program Dissemination via SUMEX*

Work in medical diagnosis is carried out with the cooperation of faculty and students in the University of Minnesota Medical School and St. Paul Ramsey Medical Center.

### *B. Sharing and Interactions with Other SUMEX-AIM Projects*

A year ago, conversations were begun with William Clancey at Stanford University regarding collaboration on the study of current knowledge capturing methods. We plan to develop this collaboration in the coming year.

### *C. Critique of Resource Management*

(None)

## **III. RESEARCH PLANS**

### *A. Project Goals and Plans*

Near term -- Our research objectives in the near term can be divided in three parts. First, we are committed to the design, implementation, and evaluation of Galen, as described above. We have completed an interactive front end so that physicians can directly enter patient data, and Galen's knowledge base is currently being "tuned" with the help of Dr. James Moller MD, an expert physician collaborator from the University of Minnesota Pediatric Cardiology Clinic. During the coming year, Galen's performance will be compared with that of the Diagnoser program and with expert physicians.

Our second objective consists of making extensions to the knowledge capturing strategies developed in our original work in medical diagnosis. In the near term this work will examine descriptive strategies in which experts attempt to use a formalized language to express what they know (e.g. production rules), observational strategies in which experts perform tasks designed to reveal information from which a theory of task specific expertise can be built, and intuitive strategies in which either experts behave as knowledge engineers or knowledge engineers attempt to perform as pseudo experts. In the coming year we will also be attempting to develop a program to automate the early stages of

knowledge capturing, analogous to the "prototype stage" of design referred to in software engineering.

Our third near term objective will be to investigate one of the central problems of recognition based problem solving, how to classify problems when solving them. Questions related to problem classification which we will be examining include: What patterns do experts and novices detect in a problem that allows them to classify it as an instance of a problem type that is already known? How does an expert make an initial choice of the level of abstraction to be used in solving a problem? How can an expert recover from an initial incorrect choice of levels? How can the difference between causal and prototypic modes of reasoning be modeled as differences in levels of abstraction, and how can a common model for these two types of reasoning be constructed? We will be pursuing these questions in the area of physics problem solving, as well as in medicine.

Long range -- Our long range objective is to improve the methodology of the "knowledge capturing" process that occurs in the early stages of the development of expert systems when problem decomposition and solution strategies are being specified. Several related questions of interest include: What are the performance consequences of different approaches, how can these consequences be evaluated, and what tools can assist in making the best choice? How can organizations be determined which not only perform well, but are structured so as to facilitate knowledge acquisition from human experts? In the coming year we will be exploring these questions in areas of design and management as well as in law, physics and medicine.

#### *B. Justification and Requirements for Continued SUMEX Use*

Our current model development takes advantage of the sophisticated Lisp programming environment on SUMEX. Although much current work with Galen is done using a version running on a local VAX 11/780, we continue to benefit from the interaction with other researchers facilitated by the SUMEX system. We expect to use SUMEX to allow other groups access to the Galen program. We also plan to continue use of the knowledge engineering tools available on SUMEX.

#### *C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM*

Paul Johnson is a member of the group of investigators who have recently submitted a proposal to establish a national computer network for cognitive scientists (COGNET). In addition, our current grant will permit the purchase of some single-user computers (we are currently comparing several alternative machines). SUMEX will continue to be used for collaborative activities and for program development requiring tools not available locally.

#### *D. Recommendations for Future Community and Resource Development*

(None)

### **II.A.3. Pilot Stanford Projects**

Following are descriptions of the informal pilot projects currently using the Stanford portion of the SUMEX-AIM resource, pending funding, full review, and authorization.

In addition to the progress reports presented here, abstracts for each project are submitted on a separate Scientific Subproject Form.

### II.A.3.1. CAMDA Project

#### CAMDA Project

##### CAMDA Research Staff:

Samuel Holtzman, Co-PI	Engineering-Economic Systems
Prof. Ronald A. Howard, Co-PI	Engineering-Economic Systems
Jack Breese	Engineering-Economic Systems
Dr. Emmet Lamb	School of Medicine
Dr. Robert Kessler	School of Medicine
Dr. Frank Polansky	School of Medicine

##### Associated faculty:

Prof. Edison Tse	Engineering-Economic Systems
Prof. Ross Shachter	Engineering-Economic Systems

## I. SUMMARY OF RESEARCH PROGRAM

### *A. Project Rationale*

The Computer-Aided Medical Decision Analysis (CAMDA) project is an attempt to develop intelligent medical decision systems by taking advantage of the complementary methodologies of decision analysis and artificial intelligence.

### *B. Medical Relevance and Collaboration*

The primary effort of the CAMDA project during 1983 was focused on the design and implementation of RACHEL, an intelligent decision system for infertile couples. This effort is aimed at helping physicians and patients deal with difficult choices regarding pertinent medical procedures. RACHEL is being developed in close cooperation between the department of Engineering-Economic Systems, the department of Obstetrics and Gynecology, and the department of Surgery (Urology Division), all at Stanford.

### *C. Highlights of Research Progress*

#### *C.1 Accomplishments this past year*

The CAMDA project began in the summer concentrated our efforts on three specific tasks: the development of a formal representation for uncertain decisions, the design and implementation of solution algorithms for formal decision problems, and the construction of an inferential processor specifically tailored to the process of formalizing decision problems.

Most of our research has been based on the concept of an influence diagram (Howard and Matheson, 1984) which is generalization of decision trees as a representation for decision problems. Influence diagrams (IDs) have several major features that make them attractive for use in intelligent decision systems. Technically, IDs prevent the loss of information often incurred in constructing asymmetric decision trees (Olmsted, 1984),

without suffering from the explicit exponential growth of symmetric decision trees. Furthermore, unlike decision trees, ID decision models take full advantage of probabilistic independence relations, which can have a significant impact on the simplicity of the decision model, and on the efficiency of its formal solution.

A particularly useful feature of influence diagrams which we have recently begun investigating is that fact that they can be used to represent deterministic as well as probabilistic relations between model elements. In fact, deterministic relations can be exploited to describe complex probabilistic behavior. This feature allows the construction of simple knowledge bases (composed primarily of deterministic statements) which can be used to create problem-specific probabilistic decision models.

In addition to their technical advantages IDs have been empirically shown to be intuitively appealing to decision makers (Owen, 1978), and to provide an excellent means of communication between experts in different fields. In the context of our own efforts, we have found that the physicians who are participating in the development of RACHEL have had little difficulty using IDs as a simple representation for expressing the decisions they and their patients face.

Another important feature of IDs is the fact that they are naturally constructed in a backwards, goal-directed fashion (decision trees usually lead to a forward-reasoning approach). Backward development of decision models has two important advantages for our purposes. First, it has a strong attention-focusing effect since it encourages the decision maker to first think of what he or she wants, and then about what can be done to change the world according to the expressed preferences. Decision trees usually have the opposite effect. Thus, they often lead the decision process along paths that although relevant to the decision at hand, have little effect on it. The attention-focusing effect of IDs on the decision making process tends to contribute to its efficiency. The second advantage of the goal-directed nature of IDs for the construction of intelligent decision systems is that it makes the formulation of decision problems amenable to computer-based automation as a rule-based system.

Having decided on IDs as a means to represent decision problems, we have designed and implemented several algorithms to solve well-formed influence diagrams. This effort has resulted in the development of a powerful software package which can generate optimal strategies and their certain equivalent directly from an ID. This package is beginning to be tested and augmented to make it easier to use by researchers other than its developers. In part, this package is based on the work of Olmsted (1983), and on a constructive proof by Shachter (1984) that, given certain technical features, shows that an influence diagram can always be solved in finite time.

An important feature of RACHEL is that it attempts to help its users in the development of models for their decisions. Thus, unlike most other decision analysis tools, RACHEL is designed to use domain knowledge. Therefore, a central element in the architecture of the RACHEL system is an algorithm which performs symbolic inference. Although several general-purpose inference-engines exist within our research environment, we have found it advantageous to implement our own for reasons of efficiency and compatibility. Furthermore, our inference algorithms are particularly well suited for the construction of decision-analytic models.

Finally, from the standpoint of computer implementation, we have developed a data structure which allows us to represent a wide class of multiple-entry disconnected cyclical directed graphs, where both vertices and edges can be associated with arbitrary data structures (such as frames). For short, we refer to these graphs as WEBS (as in a spider's), and we have used them to represent a multitude of small and medium-sized

objects such as influence diagrams, medical decision knowledge bases, command parse tables, help text databases, and mathematical data (e.g., vectors and matrices).

### *C.2 Research in progress*

The immediate goal of the CAMDA project is to complete a pilot-level implementation of the RACHEL system within the next few months. As we define it, a pilot system is one where the essential algorithms work both individually and interactively with one another, operating with knowledge that is representative of the system's domain. Such a system lacks two important elements that must exist within a prototype-level implementation: an extensive knowledge base, and a front end usable by trained users who may not be familiar with the details of the system.

To complete a pilot implementation of RACHEL, we intend to direct our efforts towards the following four tasks: incorporating a medical value model elicitation facility, strengthening our influence-diagram solution procedure, improving the performance of RACHEL's inference engine, and implementing an explanation module to justify the decision model being developed. Once this implementation is completed, RACHEL will be brought to the participating physicians to begin to develop its knowledge base.

### *D. Publications*

1. Breese, J.S., Davis, D., Parnell, G.S., and Taneja, R.: *"IDEAS: Influence Diagram Elicitation System"*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.
2. Holtzman, S.: *"A Model of the Decision Analysis Process"*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1981.
3. Holtzman, S.: *"A Decision Aid for Patients with End-Stage Renal Disease"*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.
4. Holtzman, S.: *"On the Use of Formal Models in Decision Making"*, Proc. TIMS/ORSA Joint Nat. Mtg., San Francisco, May, 1984.
5. (\*) Holtzman, S.: *"Intelligent Decision Systems"*, Ph.D. Dissertation, Department of Engineering-Economic Systems, Stanford University, forthcoming.
6. Howard, R.A., and Matheson, J.E.: *"Influence Diagrams"*, in Howard, R.A., and Matheson, J.E. (Eds.): *"The Principles and Applications of Decision Analysis,"* Vol. II, Strategic Decisions Group, Menlo Park, California, 1984.
7. Olmsted, S.M.: *"On Representing and Solving Decision Problems"*, Ph.D. Dissertation, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.
8. Owen, D.L.: *"The Use of Influence Diagrams in Structuring Complex Decision Problems"*, in Howard, R.A., and Matheson, J.E. (Eds.): *"The Principles and Applications of Decision Analysis,"* Vol. II, Strategic Decisions Group, Menlo Park, California, 1984.
9. Shachter, R.: *"Evaluating Influence Diagrams"*, working paper, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1984.

### ***E. Funding Support***

The CAMDA project does not yet have direct funding support. However, in addition to SUMEX computer usage, the project has benefited from a number of hardware gifts and research support for individuals.

#### ***E.1 Stanford Medical School***

The department of Obstetrics and Gynecology and the department of Surgery (Urology Division) have provided various types of support to the project. Samuel Holtzman has received research assistantship awards for several quarters. In addition, the Infertility Clinic at Stanford has purchased several terminals for the specific purpose of developing RACHEL and other CAMDA decision systems.

#### ***E.2 Decision Systems Laboratory***

The CAMDA project has access to the facilities of the Decision Systems Laboratory (DSL) in the Department of Engineering-Economic Systems, and constitutes the laboratory's most active research project. The DSL maintains several terminals, printers and a personal computer for research on the development of computer-based decision systems. The majority of the terminals and printers were recently donated to the DSL by Qume Corporation. MAD Computer of Santa Clara has also contributed to the support of the CAMDA project through the consignment of a MAD-1 personal computer, and provision of a research assistantship for Samuel Holtzman.

## **II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE**

### ***II.A Medical Collaborations and Program Dissemination Via SUMEX***

Since its inception, the CAMDA project has benefited from an active relationship between decision analysts, computer scientists, and members of the Stanford medical community. In particular, RACHEL is being developed in close cooperation with physicians in the Infertility Clinic at Stanford. Most of this cooperation has, up to this point, consisted of an intense mutual learning experience for all project participants. The primary purpose of this initial effort has been to develop an effective means to represent medical decision knowledge. As we have described above, this work has culminated in the definition of a representation language based on influence diagrams.

Within the next few months, RACHEL is expected to attain pilot-level performance, and its knowledge base will begin to be developed. At this point, most of the interaction involving participating physicians will shift to the design and implementation of an infertility decision knowledge base. This task will involve considerable direct use of the SUMEX facility by medical personnel.

As an added benefit of the development of RACHEL, it often occurs that specific subsystems become useful in their own right. For instance, a simple program to aid physicians in determining a course of action in cases of idiopathic infertility has been implemented and made available on SUMEX to the staff of the Stanford infertility clinic, who have used it on an experimental basis.

### **II.B. Sharing and Interactions with other SUMEX-AIM Projects**

#### ***II.B.1 SUMEX-AIM 1983 Workshop:***

Samuel Holtzman chaired the working group on decision analysis and artificial intelligence in medicine. This group considered the current status and future of medical

decision systems. A full report of the working group's deliberations and conclusions is available online at SUMEX in file <HOLTZMAN.CAMDA>AIM-DA-FINAL-REPORT.TXT, and should appear in the forthcoming workshop report.

### *II.B.2 Participation in the Knowledge Representation seminar at Stanford*

As part of the CAMDA project, we have made several presentations to the general Stanford medical and computer science community. These presentations have been made within the context of the Knowledge Representation seminar, held jointly by the computer science department and the medical school, and well attended by other SUMEX researchers at Stanford.

The speakers, and titles of the most recent presentations follow:

Samuel Holtzman:	On the Design and Implementation of Computer-Based Decision Systems.
Samuel Holtzman:	A Simple Representation for Uncertain Knowledge
Prof. Ross Shachter:	Influence Diagrams and their Use in Representing and Solving Complex Decision Problems.

### *II.C. Critique of Resource Management*

The CAMDA project has been immeasurably aided by the availability of the SUMEX computing resources. In general we find the overall physical facilities to be of excellent quality. In addition, we have been quite impressed with the quality of the SUMEX staff. In particular, we have found it to be a pleasure to deal with Ed Pattermann, who has been invariably courteous, responsive to our needs, and effective in his actions. Pam Ryalls has also provided much needed help in managing the CAMDA project in a manner that is friendly and efficient.

There are, however, two areas where we feel service and performance could be improved to the benefit of the entire SUMEX community. The first concerns the SUMEX facilities themselves, the other refers to our means of communicating with these facilities.

#### *II.C.1 SUMEX load*

In the period that the CAMDA project has been active, we have noticed a significant increase in the maximum machine loading, particularly during weekday afternoons. Although this is a normal feature of time-shared computer systems, the load has become sufficiently high in recent months, that it is beginning to be difficult to work on SUMEX during business hours. In addition, reliability has been adversely affected in some instances. An increase in SUMEX computing capacity, or a means of preventing overloading of the machine should be considered. We believe that an emphasis on distributing some of SUMEX's computing power away from a centralized mainframe could have a significant effect on reducing the system load.

#### *II.C.2 Ethernet*

The CAMDA project uses SUMEX almost exclusively through Ethernet software and hardware located in Terman Engineering Center, where the department of Engineering-Economic Systems is located. This software has on occasion been extremely unreliable for extended periods of time, resulting in substantially reduced productivity for the project. Adequate communication facilities at Stanford are of critical importance to

the successful conduct of our research. Although Stanford Ethernet management is not directly under the jurisdiction of SUMEX, in order for the SUMEX resource to be utilized to the fullest, the planning and administration of networking at Stanford needs to be better coordinated. We have begun to explore several means to improve the current situation, and we believe that explicit SUMEX support of our efforts would be quite beneficial.

### III. RESEARCH PLANS

#### *III.A Project Goals and Plans*

For the near term future the primary goals of the CAMDA project are to develop a "pilot" and then "prototype" version of the RACHEL system. Over an extended period, our objective is to arrive at useable, fully-validated and documented systems for support of medical decision making in infertility and other domains.

Implementation of a pilot system is primarily an integrative task at this point, bringing together the medical knowledge base, symbolic inference procedure, decision problem solution procedure, and influence diagram data structures. All of these components exist independently. The pilot system will consist of these systems interacting to provide a simplified version of infertility decision counseling.

The prototype implementation of RACHEL will include substantially greater amounts of medical knowledge than the pilot. The major task at this stage will be the incorporation of expert knowledge regarding functional relations, probability distributions, and decision alternatives in the infertility domain. At this point in its development, the system will be available for use by participating physicians at the infertility clinic on a "test" basis, beginning the critical phase of validation and justification of the system.

A major goal of the project is to bring RACHEL to a "defensible" level of performance in the infertility domain. A working system with full documentation, explanation of its conclusions, and user interface is envisioned. Over the long term, infertility is but a single example of the range of medical decisions amenable to decision analytic treatment in an automated system. After RACHEL has been fully implemented and tested, other systems focusing on cardiology or oncology, for example, might be developed. These systems would consist of a common core of procedural knowledge based on decision analysis, and be instantiated with the medical knowledge of the particular domain.

#### *III.B Justification and Requirements for Continued SUMEX Use*

The CAMDA project is truly interdisciplinary. It draws on elements of decision analysis, artificial intelligence, and medical science. The project has the potential to contribute to each of these disciplines in important ways. SUMEX-AIM provides the resources to continue this research with the necessary access to members of the Stanford research community.

The development of automated decision systems has the potential to greatly increase the use and acceptance of decision analysis methods. In the past, although decision analysis has been shown to be an extremely effective means of assisting in decision making in complex and uncertain domains, the cost and effort involved in producing an analysis was prohibitive for most individuals. Automated decision analysis can result in a much lower cost per user, allowing decision theoretic techniques to achieve much wider application.

The development of decision systems owes much to advancements in the fields of

artificial intelligence, expert systems, and knowledge engineering. One continuing challenge in these fields has been representation and reasoning with probabilistic knowledge. The representation of knowledge in influence diagrams, and the use of decision analysis in probabilistic reasoning are both significant topics of research being pursued within the CAMDA project.

For the medical community the CAMDA project has the potential for providing tools and techniques that greatly improve the quality of decision making in medicine. RACHEL explicitly considers uncertainty, decision alternatives, and patient preferences in developing recommendations. The objective is to develop insight and understanding regarding tradeoffs and alternatives, both for the patient and the attending physician.

SUMEX-AIM provides a unique resource for the continuation of the CAMDA project. The available computing resources, plus access to the Stanford AI and medical communities are of critical importance for the successful completion of the research.

### *III.C Needs and Plans for other Computing Resources beyond SUMEX-AIM*

We are pursuing the purchase or donation of several computing resources for installation in the Decisions System Laboratory. Our primary need at present is for a LISP machine (e.g., Symbolics 3600), enabling us to perform local processing and increase our graphics capabilities.

At present the project has access to one MAD-1 personal computer (IBM-PC type). We are considering various other PC/workstation facilities to use as front ends for CAMDA products.

### *III.D Recommendations for Future Community and Resource Development*

Increases in distributed computing capabilities on the SUMEX-AIM system is a primary need at this point. As we mentioned in Section II.C.1, distributed file editing and graphics capabilities would simultaneously reduce load on the mainframe. At this time, we are particularly interested in the possibility of designing an environment where a centralized processor (such as the SUMEX 20/60) would interact at a high level with a much less powerful dedicated processor (such as a SUN workstation, or an Apple Lisa or Macintosh) with specific capabilities such as bit-mapped graphics and special purpose hardware.

## II.A.3.2. MENTOR Project

### MENTOR Project

Stuart M. Speedie, Ph.D.  
Terrence F. Blaschke, M.D.  
Department of Medicine  
Division of Clinical Pharmacology  
Stanford University

## I. SUMMARY OF RESEARCH PROGRAM

### *A. Project Rationale*

The goal of the MENTOR (Medical EvaluationN of Therapeutic ORders) project is to design and develop an expert system for monitoring drug therapy for hospitalized patients that will provide appropriate advice to physicians concerning the existence and management of adverse drug reactions. The computer as a recording-keeping device is becoming increasingly common in hospital-based health care, but much of its potential remains unrealized. Furthermore, this information is provided to the physician in the form of raw data which is often difficult to interpret. The wealth of raw data may effectively hide important information about the patient from the physician. This is particularly true with respect to adverse reactions to drugs which can only be detected by simultaneous examinations of several different types of data including drug data, laboratory tests and clinical signs.

In order to detect and appropriately manage adverse drug reactions, sophisticated medical knowledge and problem solving is required. Expert systems offer the possibility of embedding this expertise in a computer system. Such a system could automatically gather the appropriate information from existing record-keeping systems and continually monitor for the occurrence of adverse drug reactions. Based on a knowledge base of relevant data, it could analyze incoming data and inform physicians when adverse reactions are likely to occur or when they have occurred. The MENTOR project is an attempt to explore the problems associated with the development and implementation of such a system and to implement a prototype of a drug monitoring system in a hospital setting.

### *B. Medical Relevance and Collaboration*

A number of independent studies have confirmed that the incidence of adverse reactions to drugs in hospitalized patients is significant and that they are for the most part preventable. Moreover, such statistics do not include instances of suboptimal drug therapy which may result in increased costs, extended length-of-stay, or ineffective therapy. Data in these areas are sparse, though medical care evaluations carried out as part of hospital quality assurance programs suggest that suboptimal therapy is common.

Other computer systems have been developed to influence physician decision making by monitoring patient data and providing feedback. However, most of these systems suffer from a significant structural shortcoming. This shortcoming involves the evaluation rules that are used to generate feedback. In all cases, these criteria consist of discrete, independent rules. Yet, medical decision making is a complex process in which many factors are interrelated. Thus attempting to represent medical decision-making as a